PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/10, 15/62, 15/90, C12Q 1/68, C07K 14/435

A2

(11) International Publication Number: WO 99/61604

(43) International Publication Date: 2 December 1999 (02.12.99)

(21) International Application Number: PCT/JP99/02683

(22) International Filing Date: 21 May 1999 (21.05.99)

(30) Priority Data:
10/141952 22 May 1998 (22.05.98) JP

(81) Designated States: CA, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
Without international search report and to be republished upon receipt of that report.

(71) Applicant *(for all designated States except US)*: JAPAN SCIENCE AND TECHNOLOGY CORPORATION [JP/JP]; 1-8, Hon-cho 4-chome, Kawaguchi-shi, Saitama 332-0012 (JP).

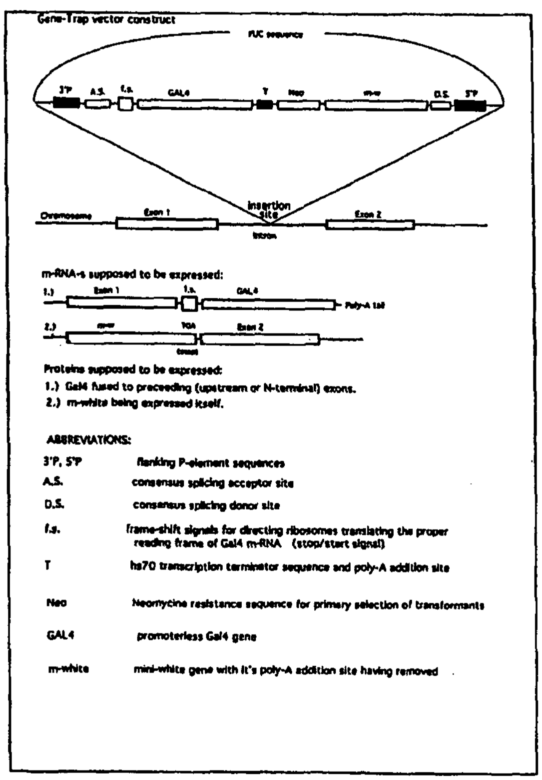
(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LUKACSOVICH, Tamas [HU/JP]; 2-30-13, Narusedai, Machida-shi, Tokyo 194-0043 (JP). ASZTALOS, Zoltan [HU/JP]; 3-16-21, Narusedai, Machida-shi, Tokyo 194-0043 (JP). YAMAMOTO, Daisuke [JP/JP]; 4-18-8, Narusedai, Machida-shi, Tokyo 194-0043 (JP). AWANO, Wakae [JP/JP]; Famiyu-102, 3-10-12, Minamidai, Sagamihara-shi, Kanagawa 228-0814 (JP).

(74) Agent: NISHIZAWA, Toshio; 6F, Mani-Building, 37-10, Udagawa-cho, Shibuya-ku, Tokyo 150-0042 (JP).

(54) Title: A VECTOR FOR GENE TRAP, AND A METHOD FOR GENE TRAPPING BY USING THE VECTOR

(57) Abstract

The present application provides a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order: an artificial consensus splicing acceptor site; a synthetic "stop/start" sequence; a reporter gene; a drug resistance gene; a gene responsible for a detectable phenotype of the *Drosophila melanogaster*, and a synthetic splicing donor site. The present application also provides a method for trapping an unknown gene of *Drosophila melanogaster* by using the vector.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

Description

A Vector for Gene Trap, and A Method for Gene Trapping by Using The Vector

Technical Field

The present invention relates to a new vector system to facilitate the cloning and functional analysis of new genes of a fly, *Drosophila melanogaster*, and a method for gene trapping with the vector system.

Background Art

There are numerous examples for application of gene trapping methods in wide range of living organisms including maize and mouse (Gossler et al,. Science, 244:463-465, 1989).

With respect to tools for gene trapping, the application of different types of enhancer trap P-element vectors (Wilson et al., Genes & Development, 3:1301-1313, 1989) for cloning and analyzing trapped genes, as well their use for mosaic analysis with the help of the Gal4/UAS transcription activator system has proven fruitful. However, sometimes the expression pattern of the Gal4 or other reporter gene of the vector construct is affected by enhancers belonging to more than one gene. Similarly, in some cases it is difficult to determine whether the enhancer trap insertion effects the function of one or more of the neighboring genes.

These circumstances altogether with the fact that in some cases the mutant phenotype could be attributed to the

changed expression of a gene with its nearest exon located more than 30 kB apart from the insertion site, can lead in unfortunate cases to an ordeal when it's time to clone and analyze the affected gene.

One object of this application is to provide a vector system that includes specifically designed artificial regulatory sequences as well as selection methods for easy screening of positive recombinant lines. More especially, this application intends to provide a vector system of this invention offering much easier and faster cloning opportunities of the affected gene, compared to the widely used enhancer trap P-element vectors. Another object of this application is to provide easier detection method possibilities of the successful trapping events and much higher chance to get more characteristic ("functional") expression patterns of the reporter gene because in the contrary with much of the cases with enhancer trap lines, when using the vector system of this invention, the reporter gene expression is influenced only by a single endogenous transcription unit and effects only the expression of the very same gene.

Disclosure of Invention

The first invention of this application is a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site.

One embodiment of the first invention is that the recombinant plasmid is derived from pCasper3.

Other embodiments of the first invention are that the reporter gene is the Gal4 gene, Gal4 DNA binding domain-P53 fusion gene or the Gal4-firefly luciferase fusion gene.

Further embodiment of this first invention is that the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene.

More further embodiment of the first invention is that the drug resistance gene is neomycin-phosphotranspherase gene and its promoter is a heatshock promoter.

The second invention of this application is a method for trapping an unknown gene of *Drosophila melanogaster* by using a vector which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site,

which method comprises the steps of:

(a) introducing the vector into the genome of a white minus fly;

(b) selecting primary transformants resistant to a drug;

(c) crossing the primary transformants with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants by picking up the flies having strong eye color,

(e) crossing the secondary transformants with UAS (Upstream Activator Sequence)-luciferase harboring strain and measuring the reporter gene expression of the resultant flies; and

(f) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

The third invention of this application is a method for trapping an unknown gene of *Drosophila melanogaster* by using a vector A which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

Gal4 DNA binding domain-P53 fusion gene as a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster;* and

a synthetic splicing donor site,

and a vector B derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs,

which method comprises the steps of:

(a) introducing each of the vectors A and B into the genomes of separate white minus flies;

(b) selecting primary transformants for the vector A which are resistant to the drug, and selecting primary transformants for the vector B which have an eye color;

(c) crossing the primary transformants for the vector A with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants for the vector A by picking up the flies having strong eye color;

(e) crossing the secondary transformants with the primary transformants for the vector B to obtain flies harboring both the vectors A and B;

(f) crossing the flies obtained in the step (e) with an UAS-luciferase harboring fly strain and measuring the reporter gene expression of the resultant flies after a heatshock treatment; and

(g) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

Embodiments of the second and third inventions are corresponded to the embodiments of the first invention, and they will be more precisely described in the following description.

Brief Description of Drawings

Figure 1 shows the schematic map of the vector of this invention, pTrap-hsneo.

Figure 2 shows the schematic map of the vector of this invention, pTrap-G4-p53.

Figure 3 shows the schematic map of the vector of this

invention, pCasperhs-G4-LT.

Figure 4 shows the schematic map of the vector of this invention, pTrap-G4-luc.

Figure 5 shows the shematic drawing of a fly genome to which the vector of this invention is inserted for cloning.

Figure 6 shows the results of sequencing RT-PCR products of aop-Gal4 and m-white-aop fusion mRNAs.

Figure 7 presents pictures of characteristic beta-galactosidase staining patterns in different parts of the fly brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

Best Mode for Carrying Out the Invention

A vector construct of the first invention, for example, can be based on the commonly used, P-element transformation vector, pCasper3 (Pirotta, Vectors: A survey of molecular cloning vectors and their uses, eds. Rodriguez, R.L. & Denhardt, D.T., Butterworths, Boston. 437-456, 1998) and the convenient Gal4-UAS expression system (Brand and Perrimon, Development, 118:401-415, 1993).

A promoterless Gal4 gene preceded by an artificial consensus splicing acceptor site and a synthetic "stop/start" sequence to govern the read through translation coming from upstream exon(s) of the trapped gene into the proper reading frame of Gal4 was inserted into the polycloning site of pCasper3.

The removal of the whole 3′ UTR (untranslated region) sequence of the mini-white gene and replacement by an artificial splicing donor site resulted in a truncated gene without its own poly-adenylation site.

Without a successful gene trapping event this truncated mini-white gene was not expected to confer any eye color, therefore in this invention a heatshock promoter directed neomycin-phosphotransferase (hs-neo) gene for helping selection of primary transformants by antibiotic feeding has been inserted.

Figure 1 shows the schematic map of the gene trap construct (pTrap-hsneo), and SEQ ID No.1 is the complete nucleotide sequence of the vector pTrap-hsneo.

Another gene trap construct, pTrap-G4-p53 (Figure 2) is created by replacing the Gal4 coding sequence of plasmid pTrap-hsneo with a Gal4 DNA binding domain-P53 fusion gene (Clontech, Matchmaker Two Hybrid System, #K1605-1). When this construct coexists in the genome of the same fly with another vector, pCasperhs-G4-LT (Figure 3) containing a heatshock promoter directed Gal4 activator domain-large T antigen (Clontech, Matchmaker Two Hybrid System, #K1605-1) fusion gene, the assembly of a functional Gal4 molecule, through p53-large T antigen interaction, can be regulated by external heatshock.

On this way, the possibility of an intentional temporary control of Gal4 activity became available. In other words the Gal4 expression in a pattern as already determined spatially by the promoter of the trapped gene now can be induced at any desired stage of development by external heatshock.

In order to make the detection of Gal4 expression easier, the Gal4 gene in another construct is replaced with a Gal4-firefly luciferase fusion gene to get pTrap-G4-luc (Figure 4). This artificial gene is coding for a fusion

polypeptide which has preserved both enzymatic activities.

The easy measuring of luciferase activity by luminoassay (Brandes et al., Neuron, 16:687-694, 1996) makes the detection of Gal4 activity comfortable in every single living fly.

Then, one of the best mode of the second or third invention, a method for gene trapping using the vector system, is described in detail.

(1) Screening:

The gene trap vector constructs can be introduced into the genome of a white minus fly by microinjection. The selection of primary transformants is possible by using G418, an analog of neomycin, resistance conferred by hs-neo gene. (When performing transformation experiments with these constructs it's turned out that the truncated mini-white gene generally provides a very slight yellow eye color which could be distinguished from w-minus phenotype in most of the cases, therefore G418 selection apparently is not necessary.)

After a line with the gene trap construct is being established, the secondary transformants can be generated on the usual way by crossing the original line with a so-called jumpstarter containing the transposase expressing delta 2-3 genetic element.

Usually a certain percentage, between 4 and 8, of the secondary transformants have much stronger eye color (deep orange or reddish) than the ancestor fly indicating that the construct was being inserted downstream of a promoter and now the mini-white gene is using the transcriptional "facilities" of that gene (e.g.: poly-adenylation site and transcriptional

terminator) instead of its removed ones. They are the most likely candidates for successful gene trap events. In case of these lines the vector probably has been inserted either into an intron of a gene or upstream from the first intron into the 5′ UTR in proper orientation (that is the direction of transcription is same for the "trapped gene" and the mini-white (and Gal4) genes as well). The mini-white gene has its own promoter therefore its expression pattern is supposed to be largely independent from that of the trapped gene.

These positive lines are to be checked in the next step for Gal4 expression by crossing them with a "marker" line harboring a UAS-luciferase reporter gene construct. (When using pTrap-G4-luc vector, this step is obviously not necessary.) Usually very strong correlation was found between eye color and Gal4 expression: more than 90% of the lines having strong eye color proved to be expressing Gal4 by means of luciferase assay using luminometer (Brandes et al., Neuron, 16:687-692, 1996).

(2) Cloning:

When the gene trap construct is being inserted into an intron of an endogenous gene, the marker genes of the construct are supposed to be spliced on mRNA level to the exons of the trapped gene by using the artificial splicing acceptor and donor sites. More exactly while the Gal4 mRNA should be joint to the exon(s) located upstream of the insertion site, at the same time the mini-white mRNA is fused to the following exon(s) accomplishing the dual tagging of the trapped gene (Figure 5).

This feature can be used for quickly and easily

identifying the trapped gene by means of 3′ and 5′ RACE (Rapid Amplification of cDNA Ends) experiments. Even cloning and sequencing only a part of the caught mRNA still provides reasonable chance to find homologous mRNAs in the BDGP (Berkeley Drosophila Genome Project) EST (Expressed Sequence Tag) library.

With these approaches, the identification of an already cloned gene can take less then a week compared to the usually more than one year period in average when analyzing a mutant created by some enhancer trap construct.

It's well-known from the literature and the present inventors also have experienced that P-element vectors tend to integrate into or near the 5′ UTR of active genes. (The present inventors found that in these cases if the insertion occurred upstream from the first intron, and therefore the artificial splicing acceptor site could not be utilized, the Gal4 gene was expressed by read-through transcription from the nearby promoter.)

The advantage of this tendency can be taken by cloning and sequencing the flanking genomic sequences of the insertion site by inverse or vectorette PCR or by plasmid rescue using suitable restriction digestion to recover the neomycin resistance gene of the construct. Then again the BDGP library can be searched to find any significant matching.

(3) Rescue:

The only reliable way to confirm that any observed mutant phenotype is really the consequence of the P-element insertion is to rescue that particular phenotype. Expectedly the phenotype (some alteration from wild type fly)is caused

by changed expression of gene(s) disturbed by insertion of the P-element. The rescue can be made by expressing the cDNA of the suspected gene most preferable with identical spatial and temporary pattern than that of the gene itself.

As it was expected, the vector constructs of the first invention usually cause strong phenotypes. It's not surprising at all because the trapped genes are supposed to be split into two parts on mRNA level resulting in null mutants in majority of the cases. Accordingly mutants obtained by this method frequently show homozygous lethality or sterility. Hypomorphic mutants can be obtained by forcing imprecise excision of the gene trap P-element construct.

As mentioned above, the Gal4 expression is obliged to reflect precisely to that of the trapped gene simply because the Gal4 gene has no its own promoter and they share a common, fused mRNA.

This identical expression provides unique opportunity to rescue the mutant phenotype by crossing this fly with another one harboring the UAS directed, cloned cDNA of the trapped gene.

On this way either the original, homozygous null mutant gene trap fly or any transheterozygous derivative of that with some hypomorphic allele over the null mutant allele can be rescued.

(4) Determination of spatial and developmental expression pattern of the trapped gene:

Histochemical determination of the spatially and temporarily controlled expression of any trapped gene is also easy following introduction of a UAS-lacZ construct into the

genome of the same fly and performing either X-gal or antibody staining for beta-galactosidase.

(5) Mosaic analysis:

Possession of a large collection of fly lines with different, characteristic and, in the case of the pTrap-G4-p53/pCasperhs-G4-TL vector system, inducible Gal4 expression pattern makes feasible carrying out mosaic analysis of virtually any gene of interest by directing the expression of their UAS-constructs on a mutant background with different Gal4 expression patterns.

This approach can answer the question of where and when that particular gene is required to be expressed to rescue the mutant phenotype.

Similarly, any gene can be expressed in different ectopic patterns to generate new dominant mutant phenotypes. This approach might help to conclude the role of that particular gene and to identify the pathway, in which it's involved.

Example

The following example illustrates a specific embodiment of the various aspects of the invention. This example is not intended to limit the invention in any manner.

Figure 6 shows the results of sequencing RT-PCR products of aop-Gal4 and m-white-aop fusion mRNAs.

The template was total RNA prepared from a positive gene trap line which has the vector pTrap-hsneo being integrated into the first intron of the well-known aop (anterior open/pokkuri/yan) developmental gene. The sequences confirm that both splicing occurred precisely at that

particular nucleotides of the artificial regulatory sequences where it was expected.

On Figure 7, there are pictures of characteristic beta-galactosidase staining patterns in different parts of the fly brain resulted from crossing positive gene trap lines with flies harboring a UAS-lacZ construct.

Industrial Applicability

The vector system of this invention offers an exceptional opportunity for easy and fast cloning of the gene responsible for the observed phenotype. Furthermore, by using the UAS-driven coding sequence of any gene of interest, that particular gene can be expressed in identical patterns than those of the trapped genes and these expressions can be regulated temporarily at any desired developmental stage.

Sequence Listing

<110> Japan Science and Technology Corporation

<120> A Vector for Gene Trap, and A Method for Gene Trapping by Using The Vector

<150> Japan, Application No. 10-141952

<151> 22 May 1998

<160> 1

<170> PatentIn Ver. 2.0

<210> 1

<211> 11206

<212> DNA

<213> Artificial sequence

<220>

<221> 3′P sequence

<222> (1)..(237)

<220>

<221> synthetic splicing acceptor site and stop/start sequence

<222> (238)..(274)

<220>

<221> Gal4 gene (coding region and 3'UTR)

<222> (275)..(3164)

<220>

<221> hsp70 terminator

<222> (3165)..(3426)

<220>

<221> synthetic junction sequence

<222> 3427-3457

<220>

<221> heat shock promoter directed neomycine resistance gene on complementer strand

<222> (3458)..(4907)

<220>

<221> mini-white gene

<222> (4908)..(8275)

<220>

<221> synthetic splicing donor site

<222> (8276)..(8299)

<220>

<221> 5'P sequence

<222> (8300)..(8446)

<220>

<221> bacterial part of pCasper3 shuttle vector including complete pUC8 sequence

<222> (8447)..(11206)

<220>

<221> synthetic DNA

<222> (238)..(274)

(3427)..(3457)

(4908)..(4914)

(8276)..(8299)

<400> 1

catgatgaaa taacataagg tggtcccgtc ggcaagagac atccacttaa cgtatgcttg 60
caataagtgc gagtgaaagg aatagtattc tgagtgtcgt attgagtctg agtgagacag 120
cgatatgatt gttgattaac ccttagcatg tccgtggggt ttgaattaac tcataatatt 180
aattagacga aattattttt aaagttttat ttttaataat ttgcgagtac gcaaagctct 240
ttctcttaca ggtcgaattg atgtgatgga tccaatgaag ctactgtctt ctatcgaaca 300
agcatgcgat atttgccgac ttaaaaagct caagtgctcc aaagaaaaac cgaagtgcgc 360
caagtgtctg aagaacaact gggagtgtcg ctactctccc aaaaccaaaa ggtctccgct 420
gactagggca catctgacag aagtggaatc aaggctagaa agactggaac agctatttct 480
actgattttt cctcgagaag accttgacat gattttgaaa atggattctt tacaggatat 540
aaaagcattg ttaacaggat tatttgtaca agataatgtg aataaagatg ccgtcacaga 600
tagattggct tcagtggaga ctgatatgcc tctaacattg agacagcata gaataagtgc 660
gacatcatca tcggaagaga gtagtaacaa aggtcaaaga cagttgactg tatcgattga 720
ctcggcagct catcatgata actccacaat tccgttggat tttatgccca gggatgctct 780
tcatggattt gattggtctg aagaggatga catgtcggat ggcttgccct tcctgaaaac 840
ggaccccaac aataatgggt tctttggcga cggttctctc ttatgtattc ttcgatctat 900
tggctttaaa ccggaaaatt acacgaactc taacgttaac aggctcccga ccatgattac 960
ggatagatac acgttggctt ctagatccac aacatcccgt ttacttcaaa gttatctcaa 1020
taattttcac ccctactgcc ctatcgtgca ctcaccgacg ctaatgatgt tgtataataa 1080
ccagattgaa atcgcgtcga aggatcaatg gcaaatcctt tttaactgca tattagccat 1140
tggagcctgg tgtatagagg gggaatctac tgatatagat gtttttttact atcaaaatgc 1200
taaatctcat ttgacgagca aggtcttcga gtcaggttcc ataattttgg tgacagccct 1260
acatcttctg tcgcgatata cacagtggag gcagaaaaca aatactagct ataattttca 1320

cagcttttcc ataagaatgg ccatatcatt gggcttgaat agggacctcc cctcgtcctt 1380
cagtgatagc agcattctgg aacaaagacg ccgaatttgg tggtctgtct actcttggga 1440
gatccaattg tccctgcttt atggtcgatc catccagctt tctcagaata caatctcctt 1500
cccttcttct gtcgacgatg tgcagcgtac cacaacaggt cccaccatat atcatggcat 1560
cattgaaaca gcaaggctct tacaagtttt cacaaaaatc tatgaactag acaaaacagt 1620
aactgcagaa aaaagtccta tatgtgcaaa aaaatgcttg atgatttgta atgagattga 1680
ggaggtttcg agacaggcac caaagttttt acaaatggat atttccacca ccgctctaac 1740
caatttgttg aaggaacacc cttggctatc ctttacaaga ttcgaactga agtggaaaca 1800
gttgtctctt atcatttatg tattaagaga ttttttcact aattttaccc agaaaaagtc 1860
acaactagaa caggatcaaa atgatcatca aagttatgaa gttaaacgat gctccatcat 1920
gttaagcgat gcagcacaaa gaactgttat gtctgtaagt agctatatgg acaatcataa 1980
tgtcaccccca tattttgcct ggaattgttc ttattacttg ttcaatgcag tcctagtacc 2040
cataaagact ctactctcaa actcaaaatc gaatgctgag aataacgaga ccgcacaatt 2100
attacaacaa attaacactg ttctgatgct attaaaaaaa ctggccactt ttaaaatcca 2160
gacttgtgaa aaatacattc aagtactgga agaggtatgt gcgccgtttc tgttatcaca 2220
gtgtgcaatc ccattaccgc atatcagtta taacaatagt aatggtagcg ccattaaaaa 2280
tattgtcggt tctgcaacta tcgcccaata ccctactctt ccggaggaaa atgtcaacaa 2340
tatcagtgtt aaatatgttt ctcctggctc agtagggcct tcacctgtgc cattgaaatc 2400
aggagcaagt ttcagtgatc tagtcaagct gttatctaac cgtccaccct ctcgtaactc 2460
tccagtgaca ataccaagaa gcacaccttc gcatcgctca gtcacgcctt ttctagggca 2520
acagcaacag ctgcaatcat tagtgccact gaccccgtct gctttgtttg gtggcgccaa 2580
ttttaatcaa agtgggaata ttgctgatag ctcattgtcc ttcactttca ctaacagtag 2640
caacggtccg aacctcataa caactcaaac aaattctcaa gcgctttcac aaccaattgc 2700
ctcctctaac gttcatgata acttcatgaa taatgaaatc acggctagta aaattgatga 2760
tggtaataat tcaaaaccac tgtcacctgg ttggacggac caaactgcgt ataacgcgtt 2820
tggaatcact acagggatgt ttaataccac tacaatggat gatgtatata actatctatt 2880
cgatgatgaa gataccccac caaacccaaa aaaagagtaa aatgaatcgt agatactgaa 2940
aaaccccgca agttcacttc aactgtgcat cgtgcaccat ctcaatttct ttcatttata 3000
catcgttttg ccttctttta tgtaactata ctcctctaag tttcaatctt ggccatgtaa 3060
cctctgatct atagaatttt ttaaatgact agaattaatg cccatctttt ttttggacct 3120

aaattcttca tgaaaatata ttacgagggc ttattcagaa gcttatcgat accgtcgact 3180
aaagccaaat agaaattatt cagttctggc ttaagttttt aaaagtgata ttatttattt 3240
ggttgtaacc aaccaaaaga atgtaaataa ctaatacata attatgttag ttttaagtta 3300
gcaacaaatt gattttagct atattagcta cttggttaat aaatagaata tatttattta 3360
aagataattc gtttttattg tcagggagtg agtttgctta aaaactcgtt tagatccact 3420
agaaggaccg cggctcctcg accggatcga aaggagggcg aagaactcca gcatgagatc 3480
cccgcgctgg aggatcatcc agccggcgtc ccggaaaacg attccgaagc ccaacctttc 3540
atagaaggcg gcggtggaat cgaaatctcg tgatggcagg ttgggcgtcg cttggtcggt 3600
catttcgaac cccagagtcc cgctcagaag aactcgtcaa gaaggcgata gaaggcgatg 3660
cgctgcgaat cgggagcggc gataccgtaa agcacgagga agcggtcagc ccattcgccg 3720
ccaagctctt cagcaatatc acgggtagcc aacgctatgt cctgatagcg gtccgccaca 3780
cccagccggc cacagtcgat gaatccagaa aagcggccat tttccaccat gatattcggc 3840
aagcaggcat cgccatgggt cacgacgaga tcctcgccgt cgggcatgcg cgccttgagc 3900
ctggcgaaca gttcggctgg cgcgagcccc tgatgctctt cgtccagatc atcctgatcg 3960
acaagaccgg cttccatccg agtacgtgct cgctcgatgc gatgtttcgc ttggtggtcg 4020
aatgggcagg tagccggatc aagcgtatgc agccgccgca ttgcatcagc catgatggat 4080
actttctcgg caggagcaag gtgagatgac aggagatcct gccccggcac ttcgcccaat 4140
agcagccagt cccttcccgc ttcagtgaca acgtcgagca cagctgcgca aggaacgccc 4200
gtcgtggcca gccacgatag ccgcgctgcc tcgtcctgca gttcattcag ggcaccggac 4260
aggtcggtct tgacaaaaag aaccgggcgc ccctgcgctg acagccggaa cacggcggca 4320
tcagagcagc cgattgtctg ttgtgcccag tcatagccga atagcctctc cacccaagcg 4380
gccggagaac ctgcgtgcaa tccatcttgt tcaatcatgc gaaacgatcc tcatcctgtc 4440
tcttgatcag atcccctatt cagagttctc ttcttgtatt caataattac ttcttggcag 4500
atttcagtag ttgcagttga tttacttggt tgctggttac ttttaattga ttcactttaa 4560
cttgcacttt actgcagatt gtttagcttg ttcagctgcg cttgtttatt tgcttagctt 4620
tcgcttagcg acgtgttcac tttgcttgtt tgaattgaat tgtcgctccg tagacgaagc 4680
gcctctattt atactccggc gctcttttcg cgaacattcg aggcgcgctc tctcgaacca 4740
acgagagcag tatgccgttt actgtgtgac agagtgagag agcattagtg cagagaggga 4800
gagacccaaa aagaaaagag agaataacga ataacggcca gagaaatttc tcgagttttc 4860
tttctgccaa acaaatgacc taccacaata accagtttgt tttgggatct agtccctaat 4920

tctagtatgt atgtaagtta ataaaaccct tttttggaga atgtagattt aaaaaaacat 4980
attttttttt tattttttac tgcactggac atcattgaac ttatctgatc agttttaaat 5040
ttacttcgat ccaagggtat ttgaagtacc aggttctttc gattacctct cactcaaaat 5100
gacattccac tcaaagtcag cgctgtttgc ctccttctct gtccacagaa atatcgccgt 5160
ctctttcgcc gctgcgtccg ctatctcttt cgccaccgtt tgtagcgtta cctagcgtca 5220
atgtccgcct tcagttgcac tttgtcagcg gtttcgtgac gaagctccaa gcggtttacg 5280
ccatcaatta aacacaaagt gctgtgccaa aactcctctc gcttcttatt tttgtttgtt 5340
ttttgagtga ttggggtggt gattggtttt gggtgggtaa gcaggggaaa gtgtgaaaaa 5400
tcccggcaat gggccaagag gatcaggagc tattaattcg cggaggcagc aaacacccat 5460
ctgccgagca tctgaacaat gtgagtagta catgtgcata catcttaagt tcacttgatc 5520
tataggaact gcgattgcaa catcaaattg tctgcggcgt gagaactgcg acccacaaaa 5580
atcccaaacc gcaatcgcac aaacaaatag tgacacgaaa cagattattc tggtagctgt 5640
gctcgctata taagacaatt tttaagatca tatcatgatc aagacatcta aaggcattca 5700
ttttcgacta cattcttttt tacaaaaaat ataacaacca gatattttaa gctgatccta 5760
gatgcacaaa aaataaataa aagtataaac ctacttcgta ggatacttcg ttttgttcgg 5820
ggttagatga gcataacgct tgtagttgat atttgagatc ccctatcatt gcagggtgac 5880
agcggacgct tcgcagagct gcattaacca gggcttcggg caggccaaaa actacggcac 5940
gctcctgcca cccagtccgc cggaggactc cggttcaggg agcggccaac tagccgagaa 6000
cctcacctat gcctggcaca atatggacat ctttggggcg gtcaatcagc cgggctccgg 6060
atggcggcag ctggtcaacc ggacacgcgg actattctgc aacgagcgac acataccggc 6120
gcccaggaaa catttgctca agaacggtga gtttctattc gcagtcggct gatctgtgtg 6180
aaatcttaat aaagggtcca attaccaatt tgaaactcag tttgcggcgt ggcctatccg 6240
ggcgaacttt tggccgtgat gggcagttcc ggtgccggaa agacgaccct gctgaatgcc 6300
cttgcctttc gatcgccgca gggcatccaa gtatcgccat ccgggatgcg actgctcaat 6360
ggccaacctg tggacgccaa ggagatgcag gccaggtgcg cctatgtcca gcaggatgac 6420
ctctttatcg gctccctaac ggccagggaa cacctgattt tccaggccat ggtgcggatg 6480
ccacgacatc tgacctatcg gcagcgagtg gcccgcgtgg atcaggtgat ccaggagctt 6540
tcgctcagca aatgtcagca cacgatcatc ggtgtgcccg gcagggtgaa aggtctgtcc 6600
ggcggagaaa ggaagcgtct ggcattcgcc tccgaggcac taaccgatcc gccgcttctg 6660
atctgcgatg agcccacctc cggactggac tcatttaccg cccacagcgt cgtccaggtg 6720

ctgaagaagc tgtcgcagaa gggcaagacc gtcatcctga ccattcatca gccgtcttcc 6780
gagctgtttg agctctttga caagatcctt ctgatggccg agggcagggt agctttcttg 6840
ggcactccca gcgaagccgt cgacttcttt tcctagtgag ttcgatgtgt ttattaaggg 6900
tatctagcat tacattacat ctcaactcct atccagcgtg ggtgcccagt gtcctaccaa 6960
ctacaatccg gcggactttt acgtacaggt gttggccgtt gtgccccggac gggagatcga 7020
gtcccgtgat cggatcgcca agatatgcga caattttgct attagcaaag tagcccggga 7080
tatggagcag ttgttggcca ccaaaaattt ggagaagcca ctggagcagc cggagaatgg 7140
gtacacctac aaggccacct ggttcatgca gttccgggcg gtcctgtggc gatcctggct 7200
gtcggtgctc aaggaaccac tcctcgtaaa agtgcgactt attcagacaa cggtgagtgg 7260
ttccagtgga aacaaatgat ataacgctta caattcttgg aaacaaattc gctagatttt 7320
agttagaatt gcctgattcc acacccttct tagttttttt caatgagatg tatagtttat 7380
agttttgcag aaaataaata aatttcattt aactcgcgaa catgttgaag atatgaatat 7440
taatgagatg cgagtaacat tttaatttgc agatggttgc catcttgatt ggcctcatct 7500
ttttgggcca acaactcacg caagtgggcg tgatgaatat caacggagcc atcttcctct 7560
tcctgaccaa catgaccttt caaaacgtct ttgccacgat aaatgtaagt cttgtttaga 7620
atacatttgc atattaataa tttactaact ttctaatgaa tcgattcgat ttaggtgttc 7680
acctcagagc tgccagtttt tatgagggag gcccgaagtc gactttatcg ctgtgacaca 7740
tactttctgg gcaaaacgat tgccgaatta ccgctttttc tcacagtgcc actggtcttc 7800
acggcgattg cctatccgat gatcggactg cgggccggag tgctgcactt cttcaactgc 7860
ctggcgctgg tcactctggt ggccaatgtg tcaacgtcct tcggatatct aatatcctgc 7920
gccagctcct cgacctcgat ggcgctgtct gtgggtccgc cggttatcat accattcctg 7980
ctctttggcg gcttcttctt gaactcgggc tcggtgccag tatacctcaa atggttgtcg 8040
tacctctcat ggttccgtta cgccaacgag ggtctgctga ttaaccaatg ggcggacgtg 8100
gagccgggcg aaattagctg cacatcgtcg aacaccacgt gccccagttc gggcaaggtc 8160
atcctggaga cgcttaactt ctccgccgcc gatctgccgc tggactacgt gggtctggcc 8220
attctcatcg tgagcttccg ggtgctcgca tatctggctc taagacttcg ggcccgacgc 8280
aaggagtaga aggtaagtag cggccgcacg taagggttaa tgttttcaaa aaaaaattcg 8340
tccgcacaca acctttcctc tcaacaagca aacgtgcact gaatttaagt gtatacttcg 8400
gtaagcttcg gctatcgacg ggaccacctt atgttatttc atcatgggcc agacccacgt 8460
agtccagcgg cagatcggcg gcggagaagt taagcgtctc caggatgacc ttgcccgaac 8520

tggggcacgt ggtgttcgac gatgtgcagc taatttcgcc cggctccacg tccgcccatt 8580
ggttaatcag cagaccctcg ttggcgtaac ggaaccatga gaggtacgac aaccatttga 8640
ggtatactgg caccgagccc gagttcaaga agaaggcgtt tttccatagg ctccgccccc 8700
ctgacgagca tcacaaaaat cgacgctcaa gtcagaggtg gcgaaacccg acaggactat 8760
aaagatacca ggcgtttccc cctggaagct ccctcgtgcg ctctcctgtt ccgaccctgc 8820
cgcttaccgg atacctgtcc gcctttctcc cttcgggaag cgtggcgctt tctcaatgct 8880
cacgctgtag gtatctcagt tcggtgtagg tcgttcgctc caagctgggc tgtgtgcacg 8940
aaccccccgt tcagcccgac cgctgcgcct tatccggtaa ctatcgtctt gagtccaacc 9000
cggtaagaca cgacttatcg ccactggcag cagccactgg taacaggatt agcagagcga 9060
ggtatgtagg cggtgctaca gagttcttga agtggtggcc taactacggc tacactagaa 9120
ggacagtatt tggtatctgc gctctgctga agccagttac cttcggaaaa agagttggta 9180
gctcttgatc cggcaaacaa accaccgctg gtagcggtgg tttttttgtt tgcaagcagc 9240
agattacgcg cagaaaaaaa ggatctcaag aagatccttt gatcttttct acggggtctg 9300
acgctcagtg gaacgaaaac tcacgttaag ggattttggt catgagatta tcaaaaagga 9360
tcttcaccta gatcctttta aattaaaaat gaagttttaa atcaatctaa agtatatatg 9420
agtaaacttg gtctgacagt taccaatgct taatcagtga ggcacctatc tcagcgatct 9480
gtctatttcg ttcatccata gttgcctgac tccccgtcgt gtagataact acgatacggg 9540
agggcttacc atctggcccc agtgctgcaa tgataccgcg agacccacgc tcaccggctc 9600
cagatttatc agcaataaac cagccagccg gaagggccga gcgcagaagt ggtcctgcaa 9660
ctttatccgc ctccatccag tctattaatt gttgccggga agctagagta agtagttcgc 9720
cagttaatag tttgcgcaac gttgttgcca ttgctacagg catcgtggtg tcacgctcgt 9780
cgtttggtat ggcttcattc agctccggtt cccaacgatc aaggcgagtt acatgatccc 9840
ccatgttgtg caaaaaagcg gttagctcct tcggtcctcc gatcgttgtc agaagtaagt 9900
tggccgcagt gttatcactc atggttatgg cagcactgca taattctctt actgtcatgc 9960
catccgtaag atgcttttct gtgactggtg agtactcaac caagtcattc tgagaatagt 10020
gtatgcggcg accgagttgc tcttgcccgg cgtcaacacg ggataatacc gcgccacata 10080
gcagaacttt aaaagtgctc atcattggaa aacgttcttc ggggcgaaaa ctctcaagga 10140
tcttaccgct gttgagatcc agttcgatgt aacccactcg tgcacccaac tgatcttcag 10200
catcttttac tttcaccagc gtttctgggt gagcaaaaac aggaaggcaa aatgccgcaa 10260
aaaagggaat aagggcgaca cggaaatgtt gaatactcat actcttcctt tttcaatatt 10320

attgaagcat ttatcagggt tattgtctca tgagcggata catatttgaa tgtatttaga 10380
aaaataaaca aataggggtt ccgcgcacat ttccccgaaa agtgccacct gacgtctaag 10440
aaaccattat tatcatgaca ttaacctata aaaataggcg tatcacgagg ccctttcgtc 10500
tcgcgcgttt cggtgatgac ggtgaaaacc tctgacacat gcagctcccg gagacggtca 10560
cagcttgtct gtaagcggat gccgggagca gacaagcccg tcagggcgcg tcagcgggtg 10620
ttggcgggtg tcggggctgg cttaactatg cggcatcaga gcagattgta ctgagagtgc 10680
accatatgcg gtgtgaaata ccgcaccgaa tcgcgcggaa ctaacgacag tcgctccaag 10740
gtcgtcgaac aaaaggtgaa tgtgttgcgg agagcgggtg ggagacagcg aaagagcaac 10800
tacgaaacgt ggtgtggtgg aggtgaatta tgaagagggc gcgcgatttg aaaagtatgt 10860
atataaaaaa tatatcccgg tgttttatgt agcgataaac gagtttttga tgtaaggtat 10920
gcaggtgtgt aagtcttttg gttagaagac aaatccaaag tctacttgtg gggatgttcg 10980
aaggggaaat acttgtattc tataggtcat atcttgtttt tattggcaca aatataatta 11040
cattagcttt ttgagggggc aataaacagt aaacacgatg gtaataatgg taaaaaaaaa 11100
aacaagcagt tatttcggat atatgtcggc tactccttgc gtcgggcccg aagtcttaga 11160
gccagatatg cgagcacccg gaagctcacg atgagaatgg ccagac 11206

CLAIMS

1. A vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site.

2. The vector of claim 1, wherein the recombinant plasmid is derived from pCasper3.

3. The vector of claim 1 or 2, wherein the reporter gene is the Gal4 gene.

4. The vector of claim 3, which has the nucleotide sequence of SEQ ID No. 1.

5. The vector of claim 1 or 2, wherein the reporter gene is Gal4 DNA binding domain-P53 fusion gene.

6. The vector of claim 1 or 2, wherein the reporter gene is the Gal4-firefly luciferase fusion gene.

7. The vector of any one of claims 1-6, wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene.

8. The vector of any one of claims 1-7, wherein the drug resistance gene is neomycin-phosphotranspherase gene and its promoter is a heatshock promoter.

9 A vector derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs.

10. A method for trapping an unknown gene of *Drosophila melanogaster* by using a vector which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster;* and

a synthetic splicing donor site,

which method comprises the steps of:

(a) introducing the vector into the genome of a white minus fly;

(b) selecting primary transformants resistant to a drug;

(c) crossing the primary transformants with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants by picking up the flies having strong eye color,

(e) crossing the secondary transformants with UAS (Upstream Activator Sequence)-luciferase harboring strain and measuring

the reporter gene expression of the resultant flies; and

(f) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

11. The method according to claim 10, wherein the recombinant plasmid is derived from pCasper3.

12. The method according to claim 10 or 11, wherein the reporter gene in the vector is the Gal4 gene, and in the step (e) the Gal4 expression is measured.

13. The method according to claim 10 or 11, wherein the reporter gene of the vector is the Gal4-firefly luciferase fusion gene, and in the step (e) expression of said fusion gene is measured without crossing the secondary transformants with UAS-luciferase harboring strain.

14. The method according to any one of claims 10 to 14, wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene, and in the step (f) the cDNAs fused to the reporter gene and the mini-white gene are cloned and sequenced.

15. The method according to any one of claims 10 to 15, wherein the drug resistance gene is neomycin-phosphotranspherase gene and its promoter is a heatshock promoter, and in the step (b) the transformants resistant to G418 is selected.

16. A method for trapping an unknown gene of *Drosophila*

melanogaster by using a vector A which is a recombinant plasmid comprising the following nucleotide sequences in this order:

an artificial consensus splicing acceptor site;

a synthetic "stop/start" sequence;

Gal4 DNA binding domain-P53 fusion gene as a reporter gene;

a drug resistance gene;

a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and

a synthetic splicing donor site,

and a vector B derived from pCasperhs, which has the heatshock promoter directed Gal4 activator domain-large T antigen fusion gene within polycloning site of the pCasperhs,

which method comprises the steps of:

(a) introducing each of the vectors A and B into the genomes of separate white minus flies;

(b) selecting primary transformants for the vector A which are resistant to a drug, and selecting primary transformants for the vector B which have an eye color;

(c) crossing the primary transformants for the vector A with a transposase source strain to force the vector to jump into other locations;

(d) selecting secondary transformants for the vector A by picking up the flies having strong eye color;

(e) crossing the secondary transformants with the primary transformants for the vector B to obtain flies harboring both the vectors A and B;

(f) crossing the flies obtained in the step (e) with an UAS-luciferase harboring fly strain and measuring the reporter gene expression of the resultant flies after a

heatshock treatment; and

(g) identifying the trapped gene by cloning and sequencing the cDNAs fused to the reporter gene and the gene responsible for a detectable phenotype of the fly.

17. The method according to claim 16, wherein the vector A is derived from pCasper3.

18. The method according to claim 16 or 17, wherein the gene responsible for a detectable phenotype of the *Drosophila melanogaster* is mini-white gene, and in the step (g) the cDNAs fused to the reporter gene and the mini-white gene are cloned and sequenced.

19. The method according to any one of claims 16 to 18, wherein the drug resistance gene is neomycin-phosphotranspherase gene and its promoter is a heatshock promoter, and in the step (b) the transformant resistant to G418 is selected.

Fig. 1

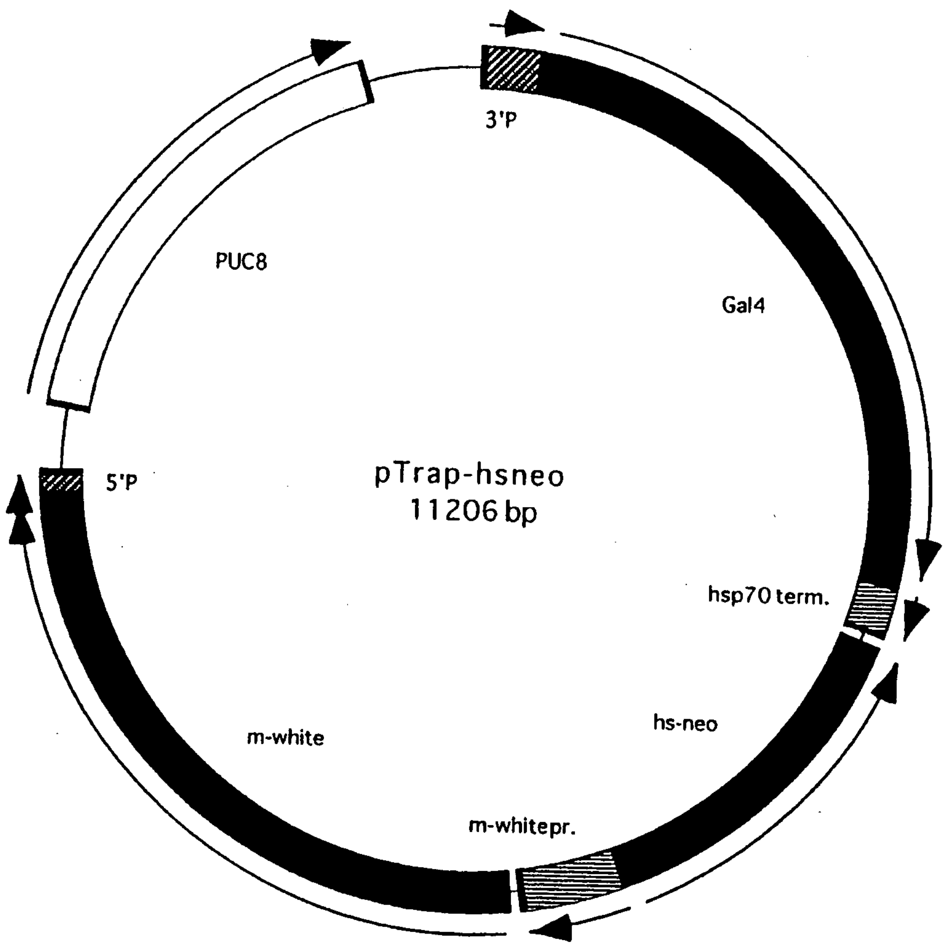

3'P
PUC8
Gal4
pTrap-hsneo
11206 bp
5'P
hsp70 term.
hs-neo
m-white
m-whitepr.

Fig. 2

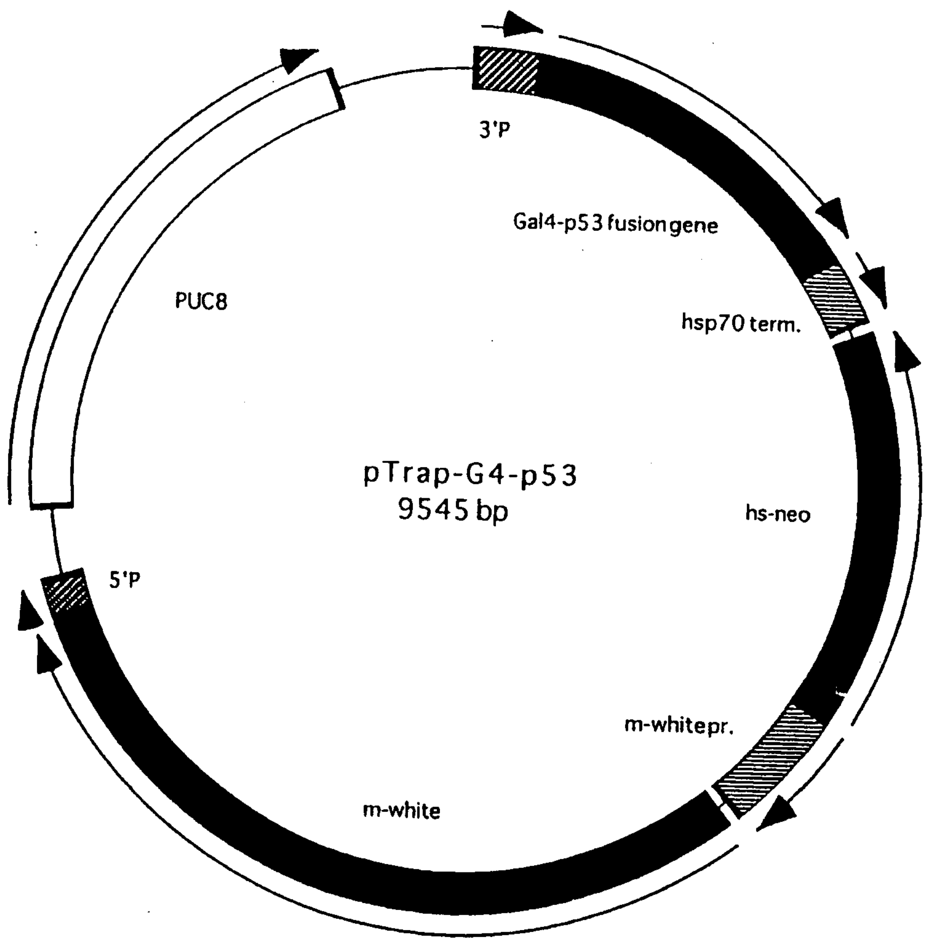

3'P
Gal4-p53 fusion gene
hsp70 term.
PUC8
pTrap-G4-p53
9545 bp
hs-neo
5'P
m-white pr.
m-white

Fig. 3

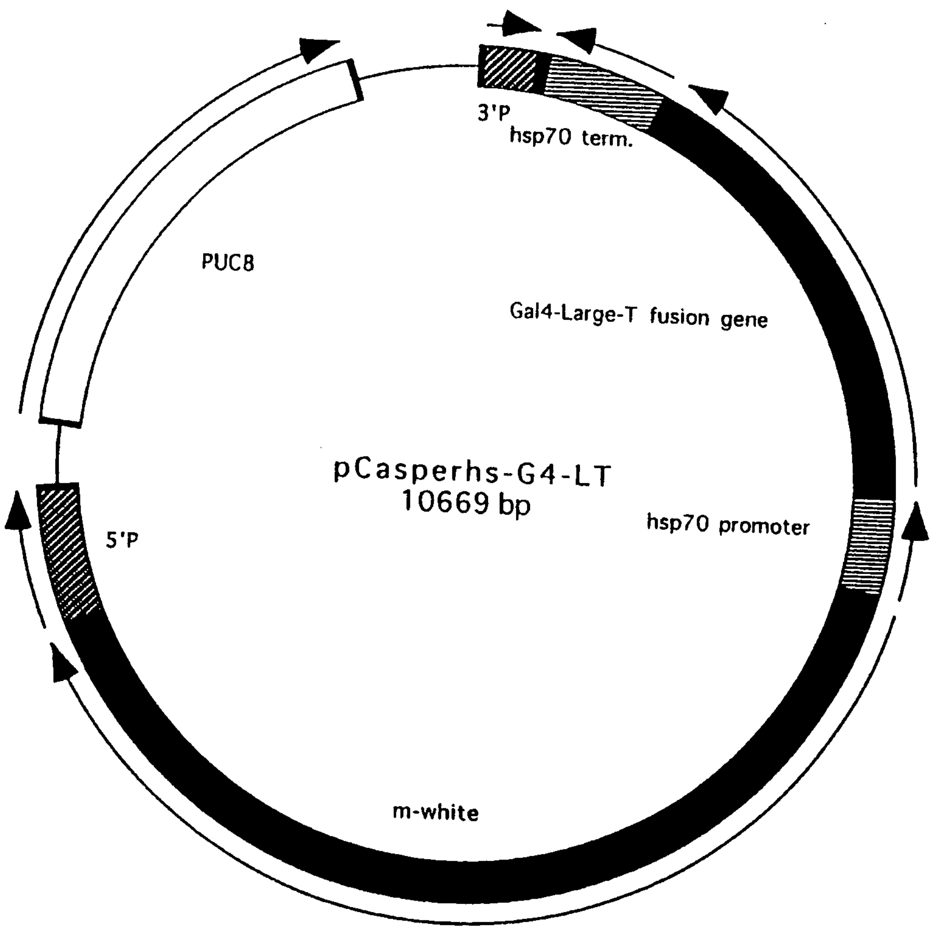

3'P
hsp70 term.
PUC8
Gal4-Large-T fusion gene
pCasperhs-G4-LT
10669 bp
hsp70 promoter
5'P
m-white

Fig. 4

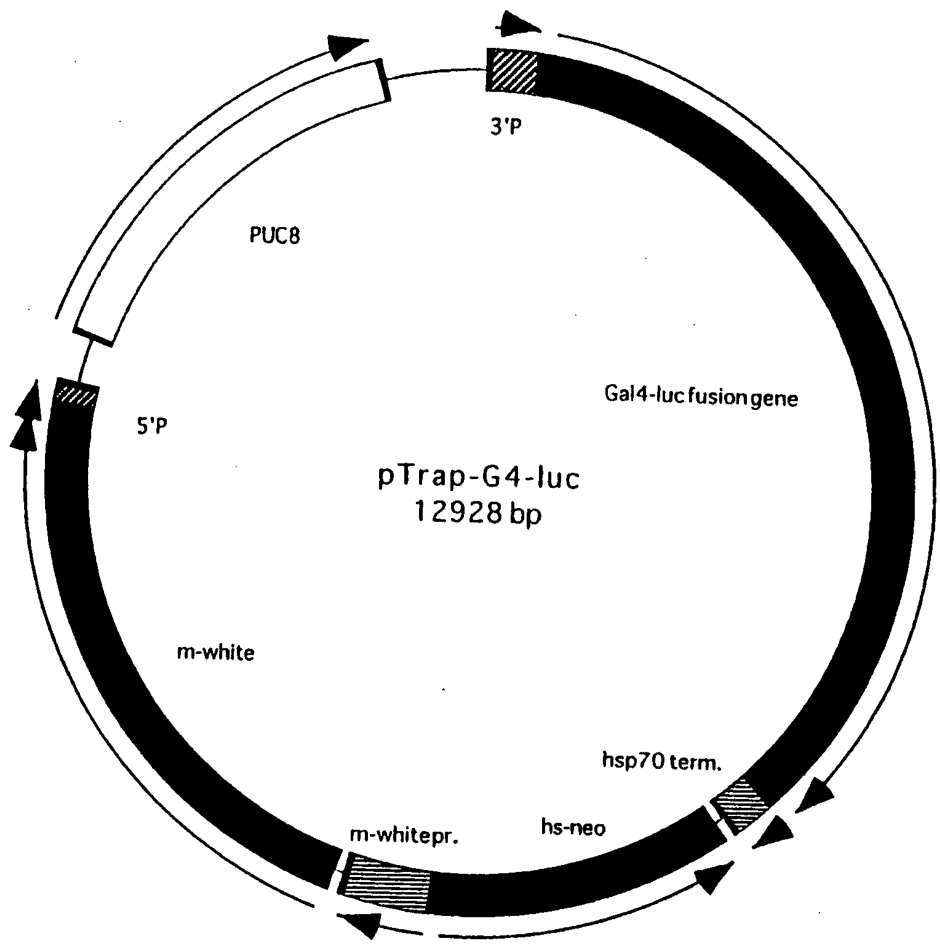
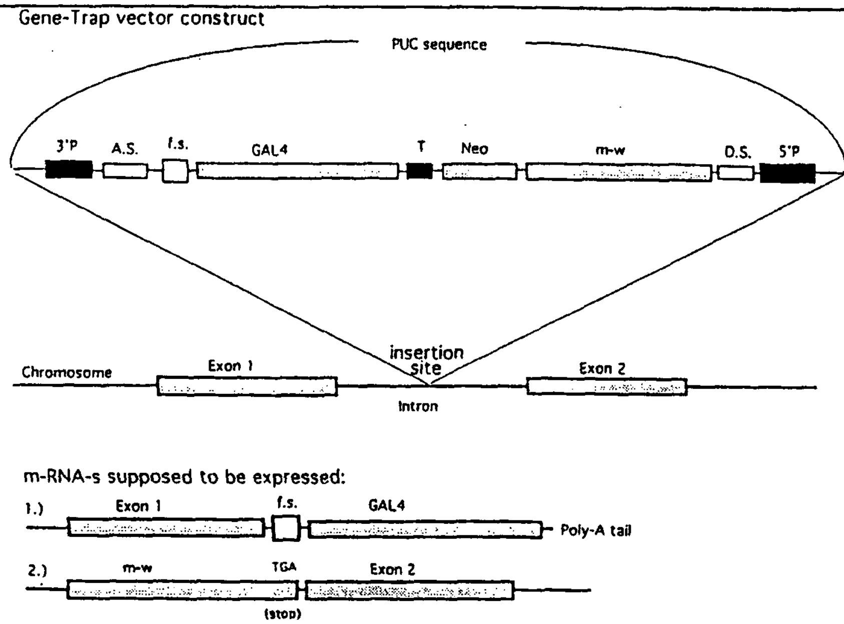

3'P
PUC8
Gal4-luc fusion gene
5'P
pTrap-G4-luc
12928 bp
m-white
hsp70 term.
hs-neo
m-white pr.

Fig. 5



Proteins supposed to be expressed:

1.) Gal4 fused to preceeding (upstream or N-terminal) exons.

2.) m-white being expressed itself.

ABBREVIATIONS:

3'P, 5'P	flanking P-element sequences
A.S.	consensus splicing acceptor site
D.S.	consensus splicing donor site
f.s.	frame-shift signals for directing ribosomes translating the proper reading frame of Gal4 m-RNA (stop/start signal)
T	hs70 transcription terminator sequence and poly-A addition site
Neo	Neomycine resistance sequence for primary selection of transformants
GAL4	promoterless Gal4 gene
m-white	mini-white gene with it's poly-A addition site having removed

Fig. 6

Precise splicing of Gal4 and mini-white genes from Gene Trap vector into anterior open gene

5'P end of vector; splice acceptor site/stop-start seq.;Gal4 gene.......mini-white gene/splice donor site; 3'P end of GT vector

.........tttgcgagtacgcaaagctctttctcttacag/gtcgaattgatgtgatggatccaatgaagct...........agacttcgggcccgacgcaaggagtagaag/gtaagtagcggccgcacgtaagggttaatg........

anterior open cDNA exon 1 exon 2

cagtacgtaaccacagtgcacagtggagcaaacag/ataacaccgcttaaaagcaattccagtggcttccac

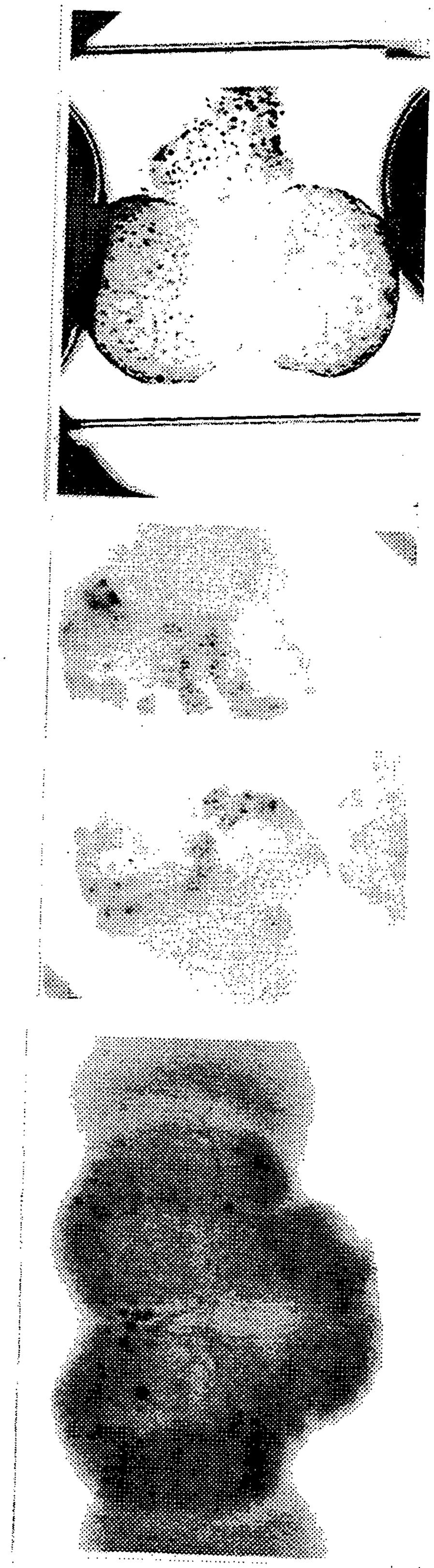
anterior open exon 1 - Gal4 fusion cDNA

cagtacgtaaccacagtgcacagtggagcaaacag/gtcgaattgatgtgatggatccaatgaagct

mini-white - anterior open exon 2 fusion cDNA

agacttcgggcccgacgcaaggagtagaag/ataacaccgcttaaaagcaattccagtggcttccac

Fig. 7



Gal4 expression patterns revealed by UAS-lacZ reporter construct.

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : **C12N 15/10, 15/62, 15/85, 15/90, C12Q 1/68, C07K 14/435**	A3	(11) International Publication Number: **WO 99/61604** (43) International Publication Date: 2 December 1999 (02.12.99)

(21) International Application Number: PCT/JP99/02683

(22) International Filing Date: 21 May 1999 (21.05.99)

(30) Priority Data:
10/141952 22 May 1998 (22.05.98) JP

(71) Applicant *(for all designated States except US)*: JAPAN SCIENCE AND TECHNOLOGY CORPORATION [JP/JP]; 1–8, Hon–cho 4–chome, Kawaguchi–shi, Saitama 332–0012 (JP).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: LUKACSOVICH, Tamas [HU/JP]; 2–30–13, Narusedai, Machida–shi, Tokyo 194–0043 (JP). ASZTALOS, Zoltan [HU/JP]; 3–16–21, Narusedai, Machida–shi, Tokyo 194–0043 (JP). YAMAMOTO, Daisuke [JP/JP]; 4–18–8, Narusedai, Machida–shi, Tokyo 194–0043 (JP). AWANO, Wakae [JP/JP]; Famiyu–102, 3–10–12, Minamidai, Sagamihara–shi, Kanagawa 228–0814 (JP).

(74) Agent: NISHIZAWA, Toshio; 6F, Mani–Building, 37–10, Udagawa–cho, Shibuya–ku, Tokyo 150–0042 (JP).

(81) Designated States: CA, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

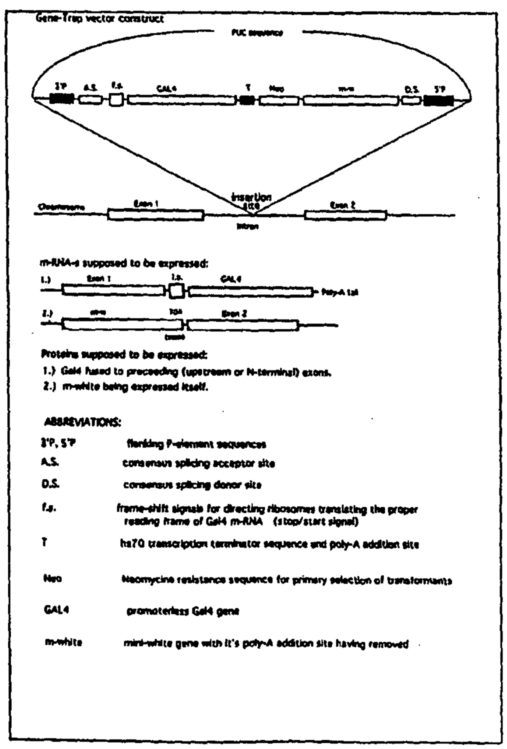
Published
With international search report.

(88) Date of publication of the international search report: 2 March 2000 (02.03.00)

(54) Title: A VECTOR FOR GENE TRAP, AND A METHOD FOR GENE TRAPPING BY USING THE VECTOR

(57) Abstract

The present application provides a vector for trapping an unknown gene of *Drosophila melanogaster*, which is a recombinant plasmid comprising the following nucleotide sequences in this order: an artificial consensus splicing acceptor site; a synthetic "stop/start" sequence; a reporter gene; a drug resistance gene; a gene responsible for a detectable phenotype of the *Drosophila melanogaster*; and a synthetic splicing donor site. The present application also provides a method for trapping an unknown gene of *Drosophila melanogaster* by using the vector.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

International Application No
PCT/JP 99/02683

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/10 C12N15/62 C12N15/85 C12N15/90 C12Q1/68
C07K14/435

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C12Q C07K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	WO 98 20031 A (JARVIK JONATHAN W) 14 May 1998 (1998-05-14) the whole document ---	1-19
A	A.H. BRAND AND N. PERRIMON: "Targeted gene expression as a means of altering cell fates and generating dominant phenotypes" DEVELOPMENT, vol. 118, 1993, pages 401-415, XP000857179 THE COMPANY OF BIOLOGISTS, LIMITED, CAMBRIDGE, GREAT BRITAIN cited in the application the whole document ---	1-19

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
3 December 1999	22/12/1999

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

1

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	W. WURST ET AL.: "A large-scale gene trap screen for insertional mutations in developmental mutations in developmentally regulated genes in mice" GENETICS, vol. 139, no. 2, February 1995 (1995-02), pages 889-899, XP000857167 GENETIC SOCIETY OF AMERICA,BALTIMORE,MD,US the whole document ---	1-19
A	C. WILSON ET AL.: "P-element-mediated enhancer detection: an efficient method for isolating and characterizing developmentally regulated genes in Drosophila" GENES & DEVELOPMENT, vol. 3, no. 9, September 1989 (1989-09), pages 1301-1313, XP000857178 CSH LABORATORY PRESS, NEW YORK,US cited in the application the whole document ---	1-19
A	P. BARTHMAIER AND E. FYRBERG: "Monitoring development and pathology of Drosophila indirect flight muscles usinf green fluorescent protein" DEVELOPMENTAL BIOLOGY, vol. 169, no. 2, June 1995 (1995-06), pages 770-774, XP002124662 ACADEMIC PRESS, INC., US the whole document ---	1-19
A	C.S. THUMMEL ET AL.: "Vectors for Drosophila P-element-mediated transformation and tissue culture transfection" GENE, vol. 74, 1988, pages 445-456, XP002124663 ELSEVIER SCIENCE PUBLISHERS,B.V.,AMSTERDAM,NL; the whole document ---	
A	V. PIRROTTA: "Vectors for P-mediated transformation in Drosophila" BIOTECHNOLOGY, VECTORS A SURVEY OF MOLECULAR CLONING VECTORS AND THEIR USES; R. L. RODRIGUEZ AND D.T. DENHARDT, vol. 1, 1988, pages 437-456, XP000857168 Butterworths, Boston, US cited in the application the whole document ---	

-/--

INTERNATIONAL SEARCH REPORT

Inte ,onal Application No
PCT/JP 99/02683

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
A	C.S. THUMMEL AND V. PIRROTTA: "New pCasPeR P element vectors" EMBL SEQUENCE DATABASE, 23 July 1996 (1996-07-23), XP002124664 Cambridge, UK Accession no. EMSYN.PEU59055; U59055; & DROS. INFO. SERVICE, vol. 71, 1992, page 150 ---	
A	A. GOSSLER ET AL.: "Mouse enbryonic stem cells and reporter constructs to detect developmentally regulated genes" SCIENCE, vol. 244, 1989, pages 463-465, XP002124665 AAAS,WASHINGTON,DC,US cited in the application the whole document -----	

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inte ional Application No
PCT/JP 99/02683

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9820031 A	14-05-1998	AU 5168598 A	29-05-1998